DAWN OF THE SEMICONDUCTOR AGE

Transistors — how we got them, what we're doing with them, and what they're doing to us

BY STEVEN E. KOMANIECKI, WB9SDN

If a scientist discovers a great revelation or unlocks some secret of the universe, few people notice — unless, of course, he runs throughout the streets like Archimedes, shouting "Eureka! I have found it." — but then they only stare. But if someone invents a new device that makes life easier for an already indolent society — well, then everyone notices.

This is what happened to semiconductor electronics. Solid-state physics was the foundation for the building of an electronic society. Its first offspring, the transistor, will ultimately change the entire dimension of human existence. Solid-state technology helped send man to the moon, deep into the seas, and into his own body; it flies his aircraft, plays his music and favorite television shows, and may even replace defective bodies. Indeed, it's a rare occasion when there's not some item within sight that's transistor related.

Because these solid-state devices influence all our lives, we should have a general understanding of them. First we should know a little about the transistor's predecessor, the vacuum tube; then we can learn how the transistor was developed, how it works, and the impact it has had on society.

Prelude to the electronic age

The foundation of the transistor was really started well before its official birthday when, in 1883, Thomas Alva Edison was striving to improve his newly invented electric light. The principal flaw in his invention was the way in which the red-hot filament slowly evaporated with sustained use. Edison tried everything possible to slow down the evaporation process. In one test, Edison took a small metal wire and placed it near the filament, thinking perhaps the presence of the wire would affect the life of the filament. It didn't; but in testing this arrangement, Edison noted an unusual phenomenon: a distinct electric current could be detected in the metal wire, even though the wire was separated from the filament by a vacuum. Somehow current had jumped from the filament to the wire. This phenomenon, though unexplained, soon came to be called the Edison effect. An example is shown in **Fig. 1**.

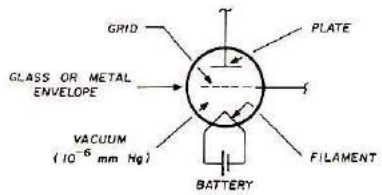
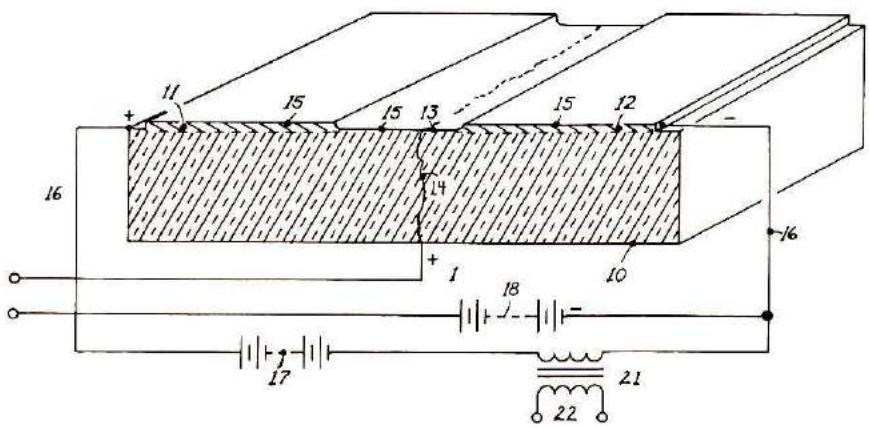


Fig. 1. Example of the Edison effect. Shown is a common triode vacuum tube (3 elements: filament, grid, and plate or anode). Electrons emanating from the heated filament cause an electric current to be induced into the nearby grid wire, even though the wire does not physically touch the filament. The phenomenon is also called "thermionic emission."

Since its discoverer found no practical use for it, the Edison effect remained unexplained and unused for more than a decade. Near the turn of the century, however, subatomic particles were discovered (Milikin, Rutherford, Chadwicke, *et al*), and the Edison effect quickly ceased to be a mystery.[1] Researchers found that all the atoms that comprise matter contain electrons (negatively charged particles) and protons (positively charged particles).[2]

An electric current was found to be caused by the flow of electrons from one atom to another throughout the length of a conductor. During this same period the physicist, Owen Williams Richardson, explained that when metal filaments are heated in a vacuum, electrons "boil" out of the filament's atoms; the electrons then leap from the filament to the solitary wire. This subatomic evaporation explained the Edison effect.

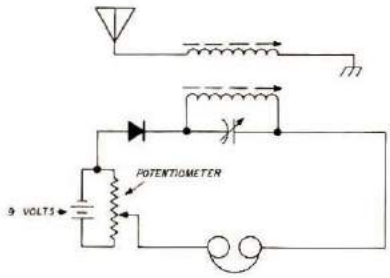


Patent drawing of Julius Lilienfield's "solid apparatus for amplifying oscillating currents," which was filed with the U.S. Patent Office in 1926. 10 designates "a base member of suitable material, for example, glass." 11 and 12 are a pair of conducting members such as a coating of platinum or silver, and 13 is an aluminum foil electrode which "is secured in position by providing a transverse fracture (14) in the glass and then re-assembling the two pieces." Finally, superimposed over 11 and 12 is a thin film of "uni-directional conductivity" (such as copper sulphide, widely used as a rectifier before the availability of silicon or germanium diodes). Lilienfield's invention is perhaps the first recorded version of the field-effect transistor.

C. David Brandle of Bell Labs in Murray Hill, New Jersey, examines a newly grown single crystal of garnet. Slices of this synthetic crystal are the basis for a new family of semiconductor memories being developed for computers (*photo courtesy Bell Labs*).

In 1904 the English electrical engineer, John Ambrose Fleming, put the Edison effect to use in an original manner. He placed a small metal cylinder around the filament and encased the entire system in an evacuated bulb. When alternating current (which changes direction many times per second) was applied to the terminals, Fleming found that current flowed mainly from the filament to the "plate" (or anode), which surrounded it. Fleming called his device a *valve* (the British still do today) because of its ability to turn the electric current on and off much like a water valve. This vacuum tube, which converts two-way motion into one-way motion, is now known as a rectifier[3] (**Fig. 2**).

The American inventor, Lee DeForest, carried electronic control a step further. In 1907, he inserted an element between the original plate and the filament. This added element was called a *grid* because it was perforated with holes; it was given a positive charge in relation to the cathode (an absence of electrons). Because it's positively charged and closer to the filament than the plate, the grid can intercept electrons



Did you ever hear of Dr. Pickard's oscillating crystal? It's one of the earliest examples of an operating semiconductor circuit, and was published in the March, 1920, issue of *QST*. The circuit was apparently widely duplicated by amateurs, but few, if any, could make it oscillate because their crystal material was not *exactly* the same as Dr. Pickard's!

from the filament, thus speeding the electron flow to the plate.

When the positive charge on the grid is made to increase slightly, the electrons come out of the filament in considerably greater quantities. A small increase or decrease in charge on the grid will impose a correspondingly large increase or decrease in the current passing from filament to plate. Furthermore, the pattern of change in current, or *signal* on the grid will be exactly duplicated by the current passing through the plate. This tube is now much more than a *valve*; it's an amplifier because of its ability to increase the amplitude or power level of a signal. **Fig. 3** shows a triode tube amplifier.

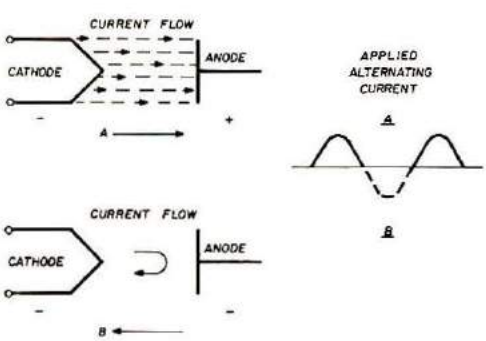


Fig. 2. Typical rectifier. Polarities (– and +) change during each half cycle of alternating current so that current can flow in direction A (sketch **A**) but not in direction B (sketch **B**).

In the United States, such amplifiers were commonly known as electron, or vacuum, tubes. Before the transistor, electron tubes were the mainstay of all electronics. Tubes were the only available means to amplify, modulate, and detect radio signals. As with all other man-made creations, the vacuum tube began to show deficiencies. The vacuum tube requires

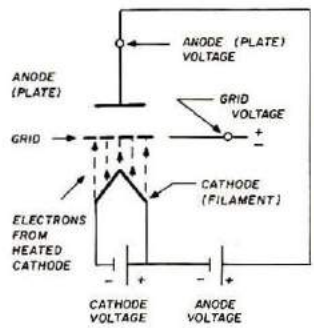


Fig. 3. A thermionic device, called a "triode" (3 elements: cathode, grid and anode) can amplify signals applied to its second element, the *grid*. The grid allows electrons from the cathode to pass through it. Variations in the grid voltage increase or decrease the electron charge from the cathode, which controls the density of the electron stream flowing between cathode (or filament) and anode. Thus, an *amplifier* is formed, which increases the applied signal.

large currents. More heat is developed, and the net effect is loss of efficiency, overheating, and shortened life span. Electron tubes also require a warm-up period. In radio circuits, tubes have an upper frequency limit beyond which they become unstable.[4] Moreover, electron tubes require a large amount of space.

Finally, because many tubes are made of glass, they lack sufficient strength to survive the rigors of field operation. Engineers and scientists recognized these limitations. They saw that no investment, manpower, or money could make the vacuum tube good enough to meet the future needs of innovations just

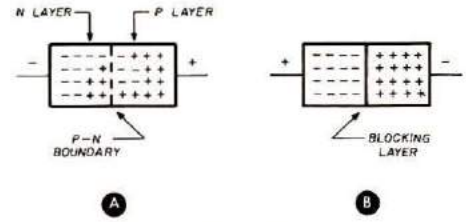


Fig. 4. The phenomenon of a semiconductor, which contains crystals contaminated by controlled addition of impurities. The type of impurity added to the crystal determines positive (P-layer) or negative (N-layer) conduction. In **A**, negative-charge carriers (N-layer) and positive-charge carriers (P-layer) migrate to the crystal center (dashed line), and current flows. If the applied voltage polarity is reversed (**B**), negative and positive charge carriers repel and migrate away from the boundary zone (blocking layer), thus no current flows.

beyond the horizon.[5] Researchers began to look in new directions: to semiconductors.

Birth of the transistor

Semiconductors were not new when transistors were first developed; in fact, they had been around since 1874 when they were first investigated by the Germans. A semiconductor is a substance with electrical properties midway between those of a conductor, such as silver or copper, and an insulator, such as porcelain. Semiconductors in one state may act as conductors, allowing current to flow, while in another state they are insulators, restricting the current flow.

Bell Telephone Laboratories became interested in these

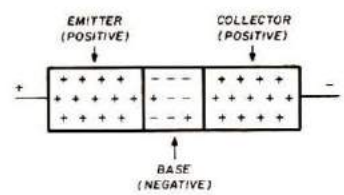
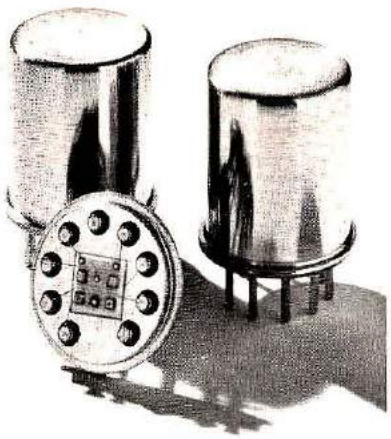
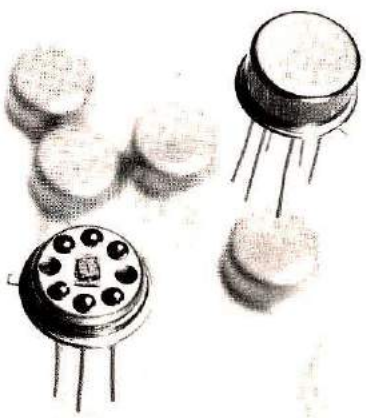


Fig. 5. How a transistor works. The current in the collector circuit is controlled by conduction in the base/emitter circuit. Thus the transistor can be used to amplify signals and to generate oscillations. Compare this action with that of "thermionic valves" or tubes (**Fig. 3**).



Field-effect transistor plug-in replacements for certain vacuum tubes, called Fetrons, are designed for use in electronic instruments (*photo courtesy Teledyne Semiconductor*).

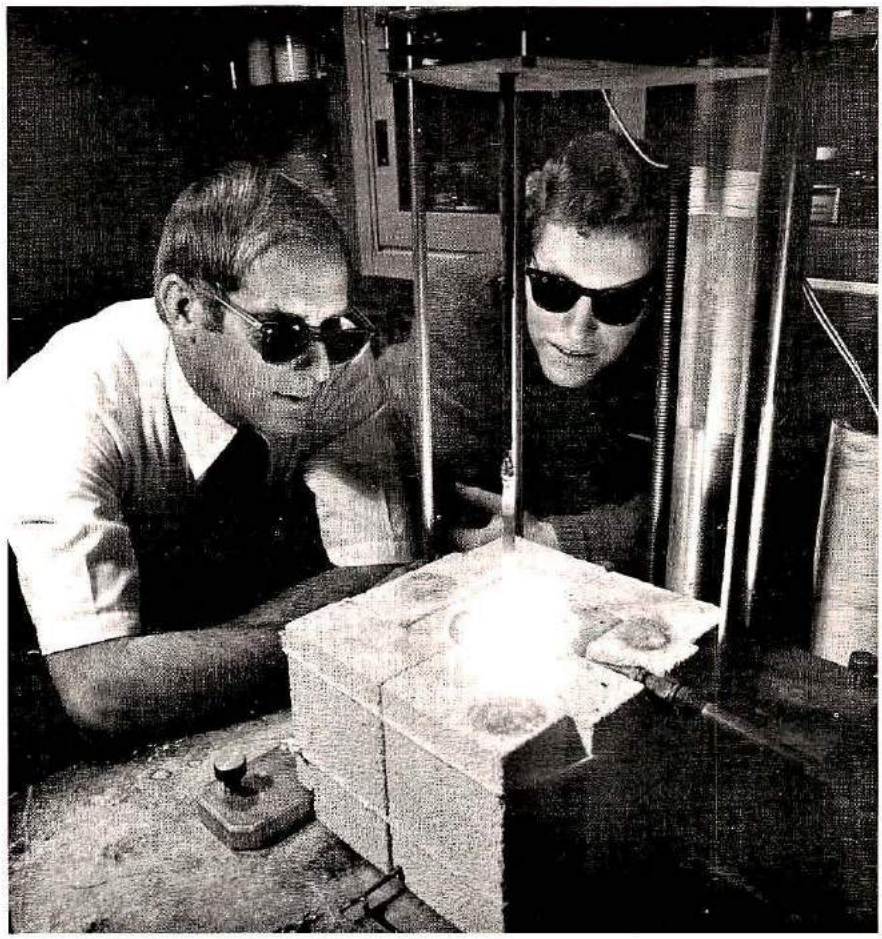
semiconductors in the early 1930s, when they sought to apply ultrahigh frequency radio waves to Bell-System communications. These frequencies were much higher than had ever been worked with before, and it was soon found that the electron tube, the conventional detector, could not detect such frequencies efficiently. Seeking another solution, the researchers thought back to the old cat-whisker detectors which had worked so well in Marconi's radio set.* The researchers found some old cat-whisker detectors in second-hand stores. The detectors made of silicon, a semi-conductor, worked best but were much too erratic to provide reliable detection. This erratic behavior later turned out to be a blessing in disguise.



Inside a typical IC. The semiconductor chip is bonded to a metal; fine gold wires are used to connect the active device to the larger leads in the header which go to the outside world. The device is then hermetically sealed with a metal cover (*photo courtesy National Semiconductor*).

The scientists melted down and pulled ingots from the purest specimens of silicon available and, in studying their behavior, noticed that there were actually two kinds of silicon: *P* and *N*, depending on the positive or negative direction in which current flowed. One day, using a different supply of silicon, the researchers accidentally produced an ingot that was not purely one type of silicon; it was *P* in one part of the block and *N* in another.

What intrigued the researchers most was the mysterious action that occurred at the *P-N* interface. Here, the dark silicon displayed an uncanny ability to convert light into electricity. A voltmeter was connected across the crystal, and when a flashlight was directed onto the silicon ingot, a definite potential difference was noted.

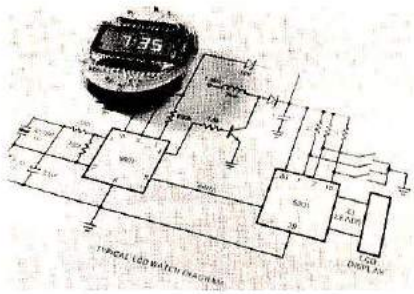


Researchers at Bell Labs watch the growth of lithium tantalate crystals which will be used in a filter for the Data Under Voice (DUV) microwave system produced in cooperation with Western Electric. Compared to quartz filters, the DUV filters have greater bandwidth and fewer components and cost less (*photo courtesy Bell Labs*).

*Cat Whisker Detector — An early form of diode detector that consists of a needle point in contact with a base material of some type of semiconductor. Used in vintage radio equipment.

Even more important, the researchers discovered that the interface acted as a rectifier, permitting current to flow in one direction only, as in a vacuum tube rectifier; no wonder the cat whiskers worked so well in Marconi's radio!

Bell Lab scientists had discovered one of the most important phenomenons of the electronic age —*P-N* junction. No one knew it at the time, but the *P-N* junction just discovered held the key not only to the transistor but also to the Bell solar cell. Several years passed, though, before these discoveries were put into practice. **Fig. 4** gives a rough idea of how the *P-N* junction works in a semiconductor device.

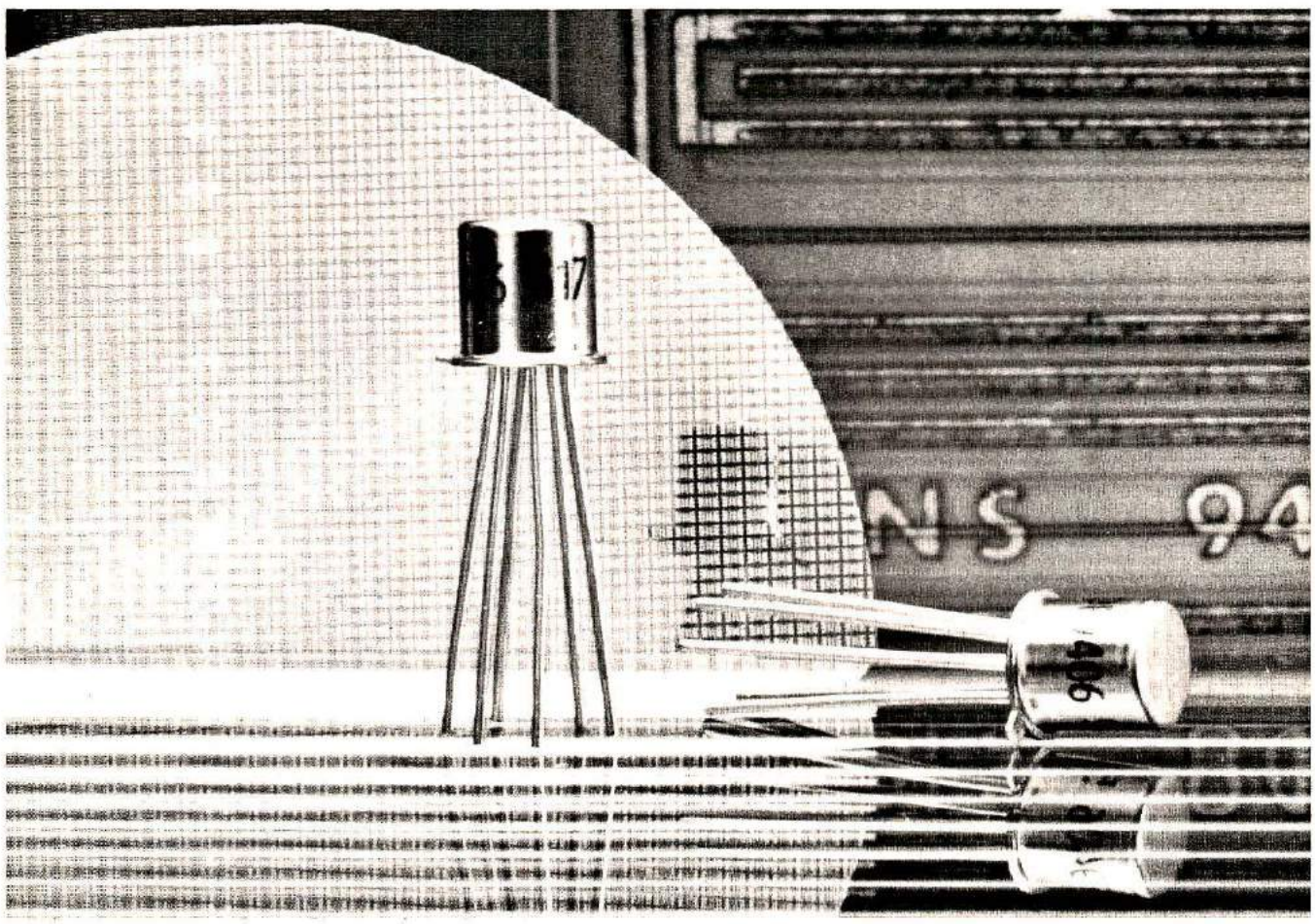


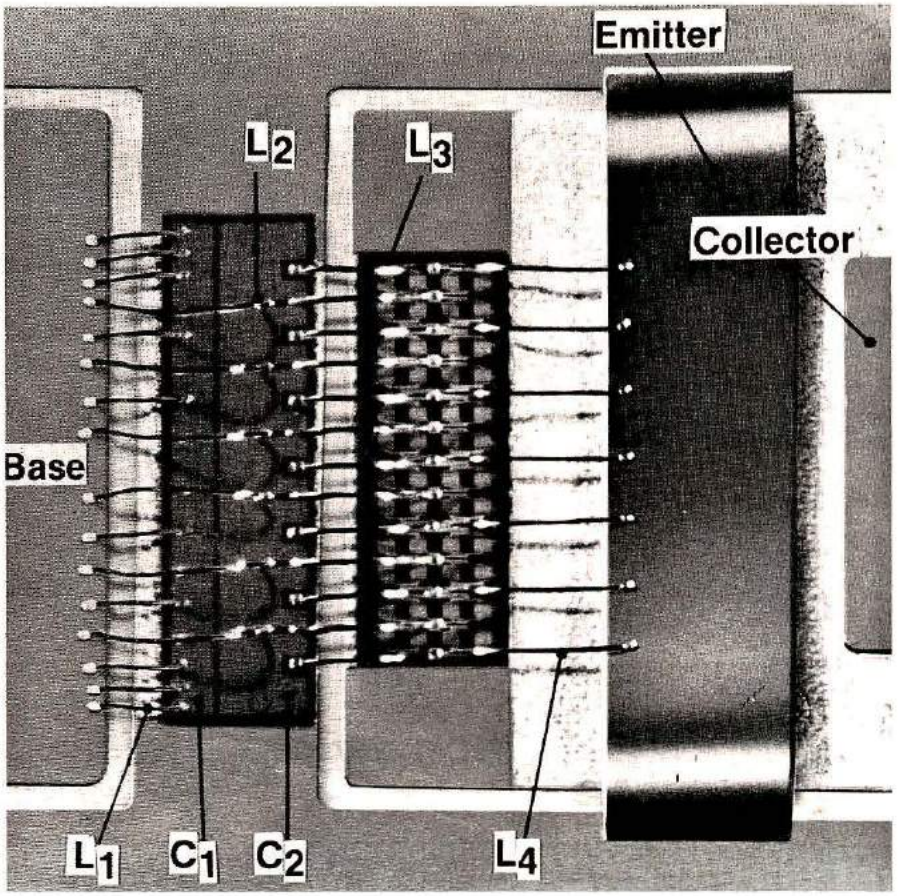
Modern digital watches would not be possible without semiconductors. This Liquid Crystal Display (LCD) watch uses two miniature ICs as well as a crystal and several external components (*photo courtesy Motorola Semiconductor*).

Meanwhile, during the 1930s, Bell engineers began to look for a substitute amplifier for the vacuum tube. Perhaps the answer could be found in semiconductor crystals. When used as detectors, the crystals involved only two electrodes, just like the filament and plate of an electron-tube rectifier. Why not introduce a third electrode that would act as a current-controller, as in the grid of a vacuum tube, to make it amplify? The answer was soon to be found.

With a return to peacetime operation after World War II, Bell Labs made an important decision: to begin an organized investigation. So in January, 1946, a research team was formed to investigate the prospect.[6] Among the many investigators who went to work was a notable trio: William Shockley, Walter H. Brattain, and John A. Bardeen. Shockley and Bardeen were pencil-and-paper physicists — theoreticians — and Brattain

Many electronic circuits require careful transistor matching for best results. The National LM194 pictured here consists of two carefully matched NPN transistors on a single chip. In the background is a microphotograph of the chip; to the left is a thin slice of silicon with hundreds of matched transistor pairs. This is cut up into single small chips which are installed in the metal packages, foreground (*photo courtesy National Semiconductor*).





Looking more like a hybrid integrated circuit than a discrete transistor, Motorola's newest line of uhf power amplifier devices features double tuned input matching networks to a multiple transistor to achieve broadband operation from 100 to 500 MHz with power outputs as high as 80 watts (*photo courtesy Motorola Semiconductor*).

was the lab man or experimenter.

Pooling their creativity, these three men evolved theories and tested them in the lab. From Shockley's calculations, Brattain tried to control the current flowing through a semiconductor crystal; no luck. Then Bardeen modified the theory and Brattain figured out a way to make it work by covering the crystal with an electrolyte. It worked, but feebly. So far, they had been using silicon crystals. Next they tried germanium, which is another semiconductor. Nearly a month later, mid December 1947, they observed that the crystal, which was in contact with two wires two-thousandths of an inch (0.05mm) apart, was behaving in a way never observed before: it was *amplifying;* it could in fact amplify a voice 40 times. This phenomenon came to be known as the *transistor effect* (see **Fig. 5**).

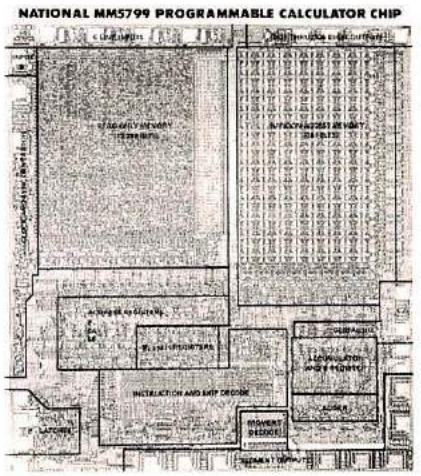
Transistor operation

A transistor is basically a resistor that amplifies electrical signals as they are transferred through it from input to output terminals. The word *transistor* was actually derived from the words *transfer* and *resistor*. It was found that transistor action depends on the activity in the *P-N* barrier.

To understand what happens inside a transistor, we must know a few important things. In a conductor (*e.g.* a wire), there are really *two* types of current carriers: electrons and something slightly different called holes. The application of an electric field to a conductor will cause electrons to move to the positive terminal and the holes (which are really vacancies of electrons originally in the metal atom) to move in the opposite direction to the negative terminal. Like charges repel, unlike charges attract.

Next we must know a little about semiconductor chemistry. The semiconductors used in electronics are usually elements that have four valence electrons each. Valence electrons are those which are free to bond with other atoms. *Silicon* and *germanium* are elements of this type. Under laboratory conditions, their crystals can be grown to nearly any desired size and can be made almost perfectly pure. This type of *intrinsic* material is not of much use, though, because it is really only a high value resistor.[7]

Soon after the first transistor creation in 1947, Bell Lab scientists started work on practical models. In 1951, Morgan Sparks created the bipolar transistor. The birth of



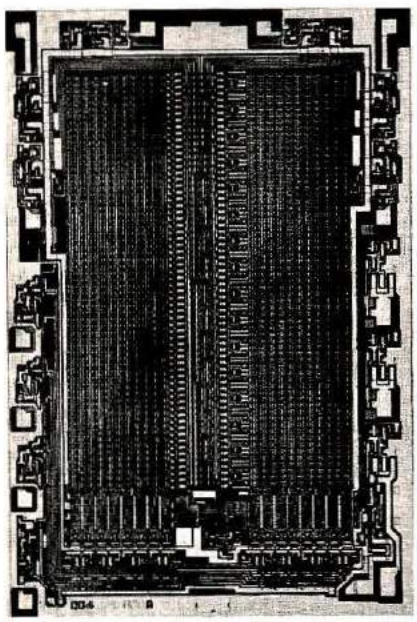
Photomicrograph of the National MM5799 programmable calculator chip, which contains all system timing, arithmetic, and logic, plus memory for more than 12000 bits. A chip such as this is called Very Large Scale Integration or VLSI (*photo courtesy National Semiconductor*).

the first point-contact transistor was announced in the laboratories' old building opposite the Hudson River docks; in the same building where, twenty-three years before, Davisson had discovered electron waves.

The auditorium was jammed with reporters, and on that day — June 30, 1948 — Shockley, Bardeen, and Brattain explained what they had done. They had samples of transistors so small that they could barely be picked up to show the curious onlookers. They also had built a working radio. It had a plastic case so that the public could see there were no tubes at all. For their work, the researchers later received the Nobel Prize in physics.

The impact of the transistor was tremendous. Why? — because Bell Telephone was determined not to keep the invention a secret. This step was radically different, because in most technical industries secrecy is the norm. Morgan Sparks, inventor of the bipolar transistor, wrote this[8] for the Bell *Record*: "Bell Lab's first important policy was not to keep transistor information secret. Not only was it not kept a secret, but we actively expounded the art as well as the science of practicing the technology. Several seminars were held in the early 1950s, where we told all we knew about transistor technology.

Nobel Prize winners Dr. William Shockley, Dr. Walter Brattain, and Dr. John Bardeen (left to right) discuss the crystal structure of semiconductors. The trio received the 1956 Nobel Physics award for their invention of the transistor, which was announced in 1948 (*photo courtesy Bell Labs*).



Photomicrograph of a Signetics 4096-bit read-only semiconductor memory. Actual size is approximately 0.1 by 0.16 inch (2.5x4mm); power dissipation is 125 microwatts per bit. This device can replace up to 144 other ICs (*photo courtesy Signetics*).

References

1. Issac Asimov, "Happy Birthday Transistor," *Saturday Review*, December 23, 1972, page 46.

2. *Understanding Transistors and Transistor Projects*, Allied Radio Corporation, Chicago, Illinois, 1969, page 9.

3. George Grammer, *Understand- Amateur Radio*, American Radio Relay League, Newington, Connecticut, 1971, page 31.

4. M. Surzberg and W. Osterheld, *Essentials of Electricity*, McGraw-Hill, New York, 1965, page 585.

5. Dr. M. Brotherton, *The Magic Crystal*, Bell Telephone, October, 1972.

6. Martin Mann, *Revolution in Electricity*, Viking Press, New York, 1962, page 62.

7. James Kyle, *Electronics Unraveled*, TAB Books, Blue Ridge Summit, Pennsylvania, 1974, page 157.

8. Morgan Sparks, "25 Years of Transistors," *Bell Laboratories Record*, page 342.

9. B. E. Strasser, "Ma Bell Led the Way," *New York Times Magazine*, December 3, 1972, page 110.

HRH